

**In the Specification:**

Please replace the paragraph (Table 1) beginning at page 94, line 1, with the following (see attached pages 94-103):

Please insert the accompanying paper copy of the Sequence Listing, page numbers 1 to 32, at the end of the application.

**REMARKS**

Applicants request entry of this amendment in adherence with 37 C.F.R. §§1.821 to 1.825. This amendment is accompanied by a floppy disk containing the above named sequences, SEQ ID NOS:1-20, in computer readable form, and a paper copy of the sequence information which has been printed from the floppy disk.

The information contained in the computer readable disk was prepared through the use of the software program "PatentIn" and is identical to that of the paper copy. This amendment contains no new matter.

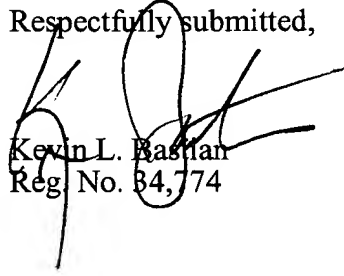
Attached hereto is a marked-up version of the changes made to the Specification and Abstract by the current Amendment. The attached pages are captioned **"VERSION WITH MARKINGS TO SHOW CHANGES MADE."**

MACK and GISH  
Application No.: 09/829,472  
Page 3

PATENT

If the Examiner believes a telephone conference would expedite prosecution of this application, please telephone the undersigned at 415-576-0200.

Respectfully submitted,



Kevin L. Bashian  
Reg. No. 34,774

TOWNSEND and TOWNSEND and CREW LLP  
Two Embarcadero Center, 8<sup>th</sup> Floor  
San Francisco, California 94111-3834  
Tel: (415) 576-0200  
Fax: (415) 576-0300  
KLB:dmw

MACK and GISH  
Application No.: 09/829,472  
Page 4



PATENT

**VERSION WITH MARKINGS TO SHOW CHANGES MADE**

**In the Specification:**

Paragraph (Table 1) beginning at line 1 of page 94 has been amended as follows (see attached pages 94-103):

SF 1281429 v1



Table 1

BCA4 DNA sequence (SEQ ID NO:1)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset Accession #: D13666; Nucleic Acid Accession #: NM\_006475; Coding sequence: 12-2522 (start and stop codons underlined)

AGAGACTCAA GATGATTCCC TTTTACCCTA TGTTTCTCT ACTATTGCTG CTTATTGTTA 60  
ACCCTATAAA CGCCAACAAT CATTATGACA AGATCTTGGC TCATAGTCGT ATCAGGGGTC 120  
GGGACCAAGG CCCAAATGTC TGTGCCCTTC AACAGATTTT GGGCACCAA AAGAAATACT 180  
TCAGCACTTG TAAGAACTGG TATAAAAAGT CCATCTGTGG ACAGAAAAAG ACTGTTTTAT 240  
ATGAATGTTG CCCTGGTTAT ATGAGAATGG AAGGAATGAA AGGCTGCCCA GCAGTTTTGC 300  
CCATTGACCA TGTATTATGGC ACTCTGGGCA TCGTGGGAGC CACCACAACG CAGCGCTATT 360  
CTGACGCCTC AAAACTGAGG GAGGAGATCG AGGGAAAGGG ATCCTTCACT TACTTTGCAC 420  
CGAGTAATGA GGCTTGGGAC AACTTGGGAT CTGATATCCG TAGAGGTTTG GAGAGCAACG 480  
TGAATGTTGA ATTACTGAAT GCTTTACATA GTCACATGAT TAATAAGAGA ATGTTGACCA 540  
AGGACTTAAA AATGGGCATG ATTATTCCTT CAATGTATAA CAATTGGGGG CTTTTCATTA 600  
ACCATTATCC TAATGGGGTT GTCACGTGTA ATTGTGCTCG AATCATCCAT GGGAAACCAGA 660  
TTGCAACAAA TGGTGTGTGC CATGTTCATT ACCGTGTGCT TACACAAATT GGTACCTCAA 720  
TTCAAGACTT CATTGAAGAG GAAGATGACC TTTTCATCTT TAGAGCAGCT GCCATCACAT 780  
CGGACATATT GGAGGCCCTT GGAAGAGACG GTCACCTCAC ACTCTTTGCT CCCACCAATG 840  
AGGCTTTTGA GAAACTTCCA CGAGGTGTCC TAGAAAGGTT CATGGGAGAC AAAGTGGCTT 900  
CCGAAGCTCT TATGAAGTAC CACATCTTAA ATACTCTCCA GTGTTCTGAG TCTATTATGG 960  
GAGGAGCAGT CTTTGAGACG CTGGAAGGAA ATACAATTGA GATAGGATGT GACGGTGACA 1020  
GTATAACAGT AAATGGAATC AAAATGGTGA ACAAAGGA TATTGTGACA AATAATGGTG 1080  
TGATCCATTG GATTGATCAG GTCCTAATTC CTGATTCTGC CAAACAAGTT ATTGAGCTGG 1140  
CTGGAAAAAC GCAACACCAC TACCGGATC TTGTGGCCCA ATTAGGCTTG GCATCTGCTC 1200  
TGAGGCCAGA TGGAGAATAC ACTTTGCTGG CACCTGTGAA TAATGCATTT TCTGATGATA 1260  
CTCTCAGCAT GGTTCAGCGC CTCCTTAAAT TAATCTGCA GAATCACATA TTGAAAGTAA 1320  
AAGTTGGCCT TAATGAGCTT TACAACGGGC AAATACTGGA AACCATCGGA GGCAACAGC 1380  
TCAGAGTCTT CGTATATCGT ACAGCTGTCT GCATTGAAAA TTCATGCATG GAGAAAGGGA 1440  
GTAAGCAAGG GAGAAACGGT GCGATTCACT TATTCCGCGA GATCATCAAG CCAGCAGAGA 1500  
AATCCCTCCA TGAAGAGTTA AAACAAGATA AGCGCTTTAG CACCTTCCTC AGCCTACTTG 1560  
AAGCTGCAGA CTTGAAAGAG CTCCTGACAC AACCTGGAGA CTGGACATTA TTTGTGCCAA 1620  
CCAATGATGC TTTTAAGGGA ATGACTAGTG AAGAAAAAGA AATTCTGATA CGGGACAAAA 1680  
ATGCTCTTCA AAACATCATT CTTTATCACC TGACACCAGG AGTTTTCATT GGAAAGGAT 1740  
TTGACCTGG TGTACTAATC ATTTTAAAGA CCACACAAGG AAGCAAAATC TTTCTGAAAG 1800  
AAGTAAATGA TACACTTCTG GTGAATGAAT TGAATCAAAA AGAATCTGAC ATCATGACAA 1860  
CAAATGGTGT AATTCATGTT GTAGATAAAC TCCTCTATCC AGCAGACACA CCTGTGGGAA 1920  
ATGATCAACT GCTGGAAATA CTTAATAAAT TAATCAAATA CATCCAAATT AAGTTTGTTT 1980  
TTGGTAGCAC CTTCAAAGAA ATCCCGTGA CTGTCTATAC AACTAAAATT ATAACCAAAG 2040  
TTGTGGAACC AAAAATTAAG GTGATTGAAG GCAGTCTTCA GCCTATTATC AAAACTGAAG 2100  
GACCCACACT AACAAAAGTC AAAATGGAAG GTGAACCTGA ATTCAGACTG ATTAAGAAG 2160  
TTGATGGAGT GCCTGTGGAA ATAACGAAA AAGAGACAG AGAAGAACGA ATCATTACAG 2220  
GTCCTGAAAT AAAATACACT AGGATTTCTA CTGGAGGTGG AGAAACAGAA GAAACTCTGA 2280  
AGAAATGTTT ACAAGAAGAG GTCACCAAGG TCACCAAAAT CATTGAAGGT GGTGATGGTC 2340  
ATTTATTGTA AGATGAAGAA ATTAAGAGAC TGCTTCAGGG AGACACACCC GTGAGGAAGT 2400  
TGCAAGCCAA CAAAAAAGTT CAAGGTTCTA GAAGACGATT AAGGGAAGGT CGTTCTCAGT 2460  
GAAAAATCAA AAACCAGAAA AAAATGTTTT TACAACCTTA AGTCAATAAC CTGACCTTAG 2520  
AAAATGTGTA GAGCCAAGTT GACTTCAGGA ACTGAAACAT CAGCACAAAG AAGCAATCAT 2580  
CAATAAATTC TGAACACAAA TTTTATTTCTG ATGAGAAACA TGAGGGAAAT 2640  
TGTGGAGTTA GCCTCTGTG GTAAAGGAAT TGAAGAAAAT ATAACACCTT ACACCTTTT 2700  
TCATCTTGAC ATTAAGAGTT CTGGCTAACT TTGGAATCCA TTAGAGAAAA ATCCTTGTC 2760  
CCGATTCAT TACAATTCAA ATCGAAGAGT TGTGAACCTG TATCCCATTTG AAAAGACCGA 2820  
GCCTTGATG TATGTTATGG ATACATAAAA TGCACGCAAG CCATTATCTC TCCATGGGAA 2880  
GCTAAGTTAT AAAAATAGGT GCTTGGTGTA CAAAACCTTT TATATCAAAA GGCTTTGCAC 2940  
ATTTCTATAT GAGTGGGTTT ACTGGTAAAT TATGTTATTT TTTACAATA ATTTTGTACT 3000  
CTCAGAATGT TTGTCAATAG CTTCTTGCAA TGCATATTTT TTAATCTCAA ACGTTTCAAT 3060  
AAAACCATTT TTCAGATATA AAGAGAATTA CTTCAAATTG AGTAATTCAG AAAAATCTCA 3120  
GATTTAAGTT AAAAAGTGGT TTGGACTTGG GAA 3180

BCA4 Protein sequence (SEQ ID NO:2)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset Accession #: D13666; Protein Accession #: NP\_006466; Predicted signal sequence: 1-21; TM domains: none; PFAM domains: fasciclin\_domains: 94-232, 234-367, 496-630; Summary: a secreted protein involved in adhesion and osteoblast development; may participate in preferential metastasis of breast cancer to bone.

MIPFLPMFSL LLLLIVNPIN ANNHYDKILA HSRIRGRDQG PNVCALQOIL GTKKKYFSTC 60  
KNWYKKSICG KQTTVLYECC PGYMRMEGMK GCPAVLPIDH VYGTGLIVGA TTTQRYSDAS 120  
KLREEIEGKG SFTYFAPSNE AWDNLDSDIR RGLSENVNVE LLNALHSHMI NKRMLTKDLK 180  
NGMIIPSMYN NLGLFINHYP NGVVTVNCAR IIHGNQIATN GVVHVIDRVL TQIGTSIQDF 240  
IEAEDDLSSF RAAATTSIDIL EALGRDGHFT LFAPTNEAFE KLPRGVLERF MGDKVASEAL 300  
MKYHILNTLQ CSESIMGGAV FETLEGNTIE ICGDGSITV NGIKMVNKKD IVTNNGVIHL 360  
IDQVLIPDSA KQVIELAGKQ QTTFTDLVAG LGLASALRPD GEYTLAPVN NAFSDDTLMS 420

VQRLKLILQ NHILKVKVGL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKGSKQG 480  
 RNGAIHIFRE IIKPAEKSLH EKLKQDKRFS TFLSLLEAAD LKELLTQPGD WTLFVPTNDA 540  
 FKGMTSEEKE ILIRDKNALQ NIILYHLTPG VFIGKGFEPG VTNILKTTQG SKIFLKEVND 600  
 TLLVNLKSK ESDIMTTNGV IHVVDKLLYP ADTPVGNDQL LEILNKLIKY IQIKFVRGST 660  
 FKEIPVTVYT TKIITKVVEP KIKVIEGSLQ PIIKTEGPTL TKVKIEGEPE FRLLIKEGETI 720  
 TEVIHGEPII KKYTKIIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKKLL 780  
 QEEVTKVTKF IEGGDGHLFE DEEIKRLLQG DTPVRKLQAN KKVQGSRRRL REGRSQ

BCA7 DNA sequence (SEQ ID NO:3)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset  
 Accession #: Z29083; Nucleic Acid Accession #: NM\_006670; Coding sequence: 85-1347 (start  
 and stop codons underlined)

CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGCGCC GTCCCAGCCC 60  
 AGCTCCGGGG AAACGCGAGC CGCGATGCGT GGGGGGTGCT CCCGGGGCCC CGCCGCCGGG 120  
 GACGGGCGTC TGCGGCTGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCCTCGTCT 180  
 TCTCCACCT CCTCGGCATC CTCCTTCTCC TCCTCGGCGC CGTTCCTGGC TTCCGCCGTG 240  
 TCCGCCACG CCCCCTGTCG GGACCACTGC CCCCCTGCTG GCGAGTGCTC CGAGGCAGCG 300  
 CGCACAGTCA AGTGCCTTAA CCGCAATCTG ACCGAGGTGC CCACGGACCT GCCCGCCTAC 360  
 GTGCGCAACC TCTTCCTTAC CGGCAACCAG CTGGCCGTGC TCCCTGCCGG CGCCTTCGCC 420  
 CGCCGGCCGC CGCTGGCGGA GCTGGCCGCG CTCACCTCA GCGGCAGCCG CCTGGACGAG 480  
 GTGCGCGCGG GCGCCTTCGA GCATCTGCCC AGCCTGCGCC AGCTCGACCT CAGCCACAAC 540  
 CCACCTGGCG ACCTCAGTCC CTTGCTTTTC TCGGCGAGCA ATGCCAGCGT CTCGGCCCCC 600  
 AGTCCCTTG TGAAGTATG CCTGAACCAC ATCGTGCCCC CTGAAGATGA GCGGCAGAAC 660  
 CGGAGCTTCG AGGCATGCT GGTGGCGGGC CTGCTGGCGG GCCGTGCACT GCAGGGGCTC 720  
 CGCGCTTGG AGCTGGCCAG CAACCACTTC CTTTACCTGC CGCGGGATGT GCTGGCCCCA 780  
 CTGCCAGCC TCAGGCACCT GGACTTAAGT AATAATTTCG TGGTGAGCCT GACCTACGTG 840  
 TCCTTCCGCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC 900  
 CTTCACAATG GCACCTTGGC TGAGTTGCAA GGTCTACCCC ACATTAGGGT TTTCTTGGAC 960  
 AACATCCCT GGGTCTGCGA CTGCCACATG GCAGACATGC TGACCTGGCT CAAGGAAACA 1020  
 GAGGTAGTGC AGGCAAGA CCGGCTCACC TGTGCATATC CGGAAAAAAT GAGGAATCGG 1080  
 GTCTCTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTCC CCCATCCCTG 1140  
 CAAACCTCTT ATGTCTTCTT GGTATTGTT TTAGCCCTGA TAGGCGCTAT TTTCTCTCTG 1200  
 GTTTTGTAAT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAACATCAG AGATGCCTGC 1260  
 AGGGATCACA TGAAGGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAACA 1320  
 AACCTCAGTT CTAACCTCGA TGCTGAGAA ATATTAGAGG ACAGACCAAG GACAACCTCTG 1380  
 CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCCACTTTCA TCCTCCACTA 1440  
 TAGATACAAC GGACTTTGAC TAAAGCAGT GAAGGGGATT TGCTTCCTTG TTATGTAAAG 1500  
 TTTCTCGGTG TGTCTGTGTA ATGTAAGACG ATGAACAGTT GTGTATAGTG TTTTACCCTC 1560  
 TTTCTTTTCT TGGAACTCCT CAACACGTAT GGAGGGATT TTCAGGTTTC AGCATGAACA 1620  
 TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC 1680  
 ACAGATAGCA TTCAACAAAA GCTGCCTCAA CTTTTTCGAG AAAAATACTT TATTATAAAA 1740  
 TATCAGTTTT ATTCTCATGT ACCTAAGTTG TGGAGAAAT AATTGCATCC TATAAATGCT 1800  
 CTGCAGAGCT TAGCAGGCTC TTCAAAATAA CTCCATGGTG CACAGGAGCA CCTGCATCCA 1860  
 AGAGCATGCT TACATTTTAC TGTTCTGCAT ATTACAAAA ATAACCTGCA ACTTCATAAC 1920  
 TTCTTTGACA AAGTAAATTA CTTTTTTGAT TGCAGTTTAT ATGAAATGT ACTGATTTTT 1980  
 TTTTAATAAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAAATAA AAAAATAAAG 2040  
 ATTCTTAAAA GAA

BCA7 Protein sequence (SEQ ID NO:4)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset  
 Accession #: Z29083; Protein Accession #: NP\_006661; Predicted Signal sequence: 1-32;  
 Predicted TM domains: 357-373; PFAM domains: leucine-rich\_repeats: 61-90, 119-142, 143-166,  
 235-258, 259-282, 294-345;  
 Summary: a type 1a TM protein of unknown function, detected in multiple cancers, with highest  
 expression in breast cancer.

MPGGCSRGA AGDGRRLRLAR LALVLLGWVS SSSPTSSASS FSSSAPFLAS AVSAQPPLPD 60  
 QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFLTQ NQLAVLPAGA FARRPPLAEL 120  
 AALNLSGSR LDEVRAFAFEH LPSLRQLDLS HNPLADLSPF AFGSGNASVS APSPLVELIL 180  
 NHIVPPEDER QNRSFEGMVV AALLAGRALQ GLRRLELASN HFLYLPDVL AQLPSLRHLD 240  
 LSNNLSVSLT YVSRNLTHL ESLHLEDNAL KVLHNGTLAE LQGLPHIRVF LDNNPWVDCD 300  
 HMADMVTWLK ETEVVQGKDR LTCAYPEKMR NRVLLELNSA DLDCDPILPP SLQTSYVFLG 360  
 IVLALIGAIF LLVLYLNRKG IKKWMHNIRD ACRDHMEGYH YRYEINADPR LTNLSNSNDV

BCX5 DNA sequence (SEQ ID NO:5)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Nucleic Acid  
 Accession #: AF160477; Coding sequence: 225-1757 (start and stop codons underlined)

GGGGAGCTCG GAGCTCCCGA TCACGGCTTC TTGGGGGTAG CTACGGCTGG GTGTGTAGAA 60  
 CGGGGCGGGG GCTGGGGCTG GGTCCCTAG TGAGACCCAA GTGCGAGAGG CAAGAACTCT 120  
 GCAGTTCCT GCCTTCTGGG TCAGTTCCTT ATTCAAGTCT GCAGCCGGCT CCCAGGGAGA 180  
 TCTCGGTGGA ACTTCAGAAA CGCTGGGCGA TCTGCCTTTC AACCATGCCC CTGTCCCTGG 240  
 GAGCCGAGAT GTGGGGGCTT GAGGCCTGGC TGCTGCTGCT GCTACTGCTG GCATCATTTA 300  
 CAGGCGGGTG CCCCCTGGGT GAGCTGGAGA CCTCAGACGT GGTAACTGTG GTGCTGGGCC 360  
 AGGACGCAAA ACTGCCTGCT TTCTACCGAG GGGACTCCGG CGAGCAAGTG GGGCAAGTGG 420

CATGGGCTCG	GGTGGACGCG	GGCGAAGGCG	CCCAGGAAC	AGCGCTACTG	CACTCCAAAT	480
ACGGGCTTCA	TGTGAGCCCG	GCTTACGAGG	GCCGCGTGGA	GCAGCCGCGG	CCCCACGCA	540
ACCCCTTGA	CGGCTCAGTG	CTCCTGCGCA	ACGAGTGCA	GGCGGATGAG	GGCGAGTACG	600
AGTGCCGGGT	CAGCCTCTTC	CCCGCCGCGA	GCTTCCAGGC	GCGGCTGCGG	CTCCGAGTGA	660
TGGTGCCTCC	CCTGCCCTCA	CTGAATCCTG	GTCCAGCACT	AGAAGAGGGC	CAGGGCCTGA	720
CCCTGGCAGC	CTCCTGCACA	GCTGAGGGCA	GCCCAGCCCC	CAGCGTGACC	TGGGACACGG	780
AGGTCAAAGG	CACAACGTCC	AGCCGTTCCT	TCAAGCACTC	CCGCTCTGCT	GCCGTCACCT	840
CAGAGTTCCA	CTTGGTGCCT	AGCCGCAGCA	TGAATGGGCA	GCCACTGACT	TGTGTGGTGT	900
CCCATCCTGG	CCTGCTCCAG	GACCAAAGGA	TCACCCACAT	CCTCCACGTG	TCCTTCCTTG	960
CTGAGGCCTC	TGTGAGGGGC	CTTGAAGACC	AAAATCTGTG	GCACATTGGC	AGAGAAGGAG	1020
CTATGCTCAA	GTGCCTGAGT	GAAGGGCAGC	CCCCCTCCCTC	ATACAACTGG	ACACGGCTGG	1080
ATGGGCCCTCT	GCCCAGTGGG	GTACGAGTGG	ATGGGGACAC	TTTGGGCTTT	CCCCACTGA	1140
CCACTGAGCA	CAGCGGCATC	TACGTCTGCC	ATGTCAGCAA	TGAGTTCTCC	TCAAGGGATT	1200
CTCAGGTCAC	TGTGGATGTT	CTTGACCCCC	AGGAAGACTC	TGGGAAGCAG	GTGGACCTAG	1260
TGTCAGCCTC	GGTGGTGGTG	GTGGGTGTGA	TCGCCGCACT	CTTGTCTGTC	CTTCTGGTGG	1320
TGGTGGTGGT	GCTCATGTCC	CGATACCATC	GGCGCAAGGC	CCAGCAGATG	ACCCAGAAAT	1380
ATGAGGAGGA	GCTGACCCCTG	ACCAGGGAGA	ACTCCATCCG	GAGGCTGCAT	TCCCATCACA	1440
CGGACCCAG	GAGCCAGCCG	GAGGAGAGTG	TAGGGCTGAG	AGCCGAGGGC	CACCCTGATA	1500
GTCTCAAGGA	CAACAGTAGC	TGCTCTGTGA	TGAGTGAAGA	GCCCGAGGGC	CGCAGTTACT	1560
CCACGCTGAC	CACGGTGAGG	GAGATAGAAA	CACAGACTGA	ACTGCTGTCT	CCAGGCTCTG	1620
GGCGGECCEA	GGAGGAGGAA	GATCAGGATG	AAGGCATCAA	ACAGGCCATG	AACCATTTTG	1680
TTCAGGAGAA	TGGGACCCTA	CGGGCCAAGC	CCACGGGCAA	TGGCATCTAC	ATCAATGGGC	1740
GGGGACACCT	GGTCTGACCC	AGGCCTGCCT	CCCTTCCCTA	GGCCTGGCTC	CTTCTGTTGA	1800
CATGGGAGAT	TTTAGTCTCAT	CTTGGGGGCC	TCCTTAAACA	CCCCCATTTT	TTGCGGAAGA	1860
TGCTCCCCAT	CCCCTGACT	GCTTGACCTT	TACCTCCAAC	CCTTCTGTTC	ATCGGGAGGG	1920
CTCCACCAAT	TGAGTCTGCT	CACCATGTGA	TGCAGGTCAC	TGTGTGTGTA	CATGTGTGCC	1980
TGTGTGAGTG	TTGACTGACT	GTGTGTGTGT	GGAGGGGTGA	CTGTCCGTGG	AGGGGTGACT	2040
GTGTCCGTGG	TGTGTATTAT	GCTGTCTAT	CAGAGTCAAG	TGAACTGTGG	TGTATGTGCC	2100
ACGGGATTTG	AGTGGTTGCG	TGGGCAACAC	TGTCAGGGTT	TGGCGTGTGT	GTCTGTGGC	2160
TGTGTGTGAC	CTCTGCCTGA	AAAAGCAGTG	ATTTTCTCAG	ACCCAGAGC	AGTATTAATG	2220
ATGCAGAGGT	TGGAGGAGAG	AGGTGGAGAC	TGTGGCTCAG	ACCCAGGTGT	GCGGCGATAG	2280
CTGGAGCTGG	AATCTGCCTC	CGGTGTGAGG	GAACCTGTCT	CCTACCACTT	CGGAGCCATG	2340
GGGGCAAGTG	TGAAGCAGCC	AGTCCCTGGG	TCAGCCAGAG	GCTTGAACCT	TTACAGAAGC	2400
CCTCTGCCCT	CTGGTGGCCT	CTGGGCCTGC	TGCATGTACA	TATTTTCTGT	AAATATACAT	2460
GCGCGGGGAG	CTTCTTGACG	GAATACTGCT	CCGAATCACT	TTTTATTTTT	TTCTTTTTTT	2520
TTTCTTGCCC	TTTCCATTAG	TTGTATTTTT	TATTTATTTT	TATTTTATT	TTTTTTTAGA	2580
GATGGAGTCT	CACATATGTTG	CTCAGGCTGG	CCTTGAACCT	CTGGGCTCAA	GCAATCCTCC	2640
TGCCTCAGCC	TCCCTAGTAG	CTGGGACTTT	AAGTGATACAC	CACGTGTCCT	GCTTTGAATC	2700
CTTTACGAAG	AGAAAAA	AATTAAGAA	AGCCTTTAGA	TTTATCCAAT	GTTTACTACT	2760
GGGATTGCTT	AAAGTGAGGC	CCCTCCAACA	CCAGGGGGTT	AATTCCTGTG	ATTGTGAAAG	2820
GGGCTACTTC	CAAGGCATCT	TCATGCAGGC	AGCCCTTGG	GAGGGCACCT	GAGAGCTGGT	2880
AGAGTCTGAA	ATTAGGGATG	TGAGCCTCGT	GGTTACTGAG	TAAGGTAAAA	TTGCATCCAC	2940
CATTGTTTGT	GATACCTTAG	GGAATTGCTT	GGACCTGGTG	ACAAGGGCTC	CTGTTCAATA	3000
GTGGTGTGTTG	GGAGAGAGAG	AGCAGTGATT	ATAGACCGAG	AGAGTAGGAG	TTGAGGTGAG	3060
GTGAAGGAGG	TGCTGGGGGT	GAGAATGTCT	CCTTCCCCC	TGGGTTTTGG	ATCACTAATT	3120
CAAGGCTCTT	CTGGATGTTT	CTCTGGGTTG	GGGCTGGAGT	TCAATGAGGT	TTATTTTATG	3180
CTGGCCCAAC	CAGATACACT	CAGCCAGAAAT	ACCTAGATT	AGTACCCAAA	CTCTCTTAG	3240
TCTGAAATCT	GCTGGATTTC	TGGCCTAAGG	GAGAGCTCC	CATCTTCGT	TCCCGAGCCA	3300
GCCTAGGACT	TGGAATGTGG	AGCCTGAAGA	TCTAAGATCC	TAACATGTAC	ATTTTATGTA	3360
AATATGTGCA	TATTTGTACA	TAAATGATA	TTCTGTTTTT	AAATAAACAG	ACAAAACCTG	3420
TTCAAAAAA	AAAAAAA	AAAAAAA				

# BCX5 Protein sequence (SEQ ID NO:6)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Protein Accession #: AF160477; Predicted Signal sequence: 1-26; Predicted TM domains: 355-371; PFAM domains: IgSF\_domain: 45-129, 162-225, 263-317; Summary: A type Ia TM protein; is a member of the immunoglobulin superfamily.

MPLSLGAEMW	GPEAWLLLLL	LLASFTGRCP	AGELETSDEV	TVVLGQDAKL	PCFYRGDSGE	60
QVGQVAVARV	DAGEGAQELA	LLHSKYGLHV	SPAYEGRVEQ	PPPPRNPLDG	SVLLRNAVQA	120
DEGEYECRV	TFPAGSFQAR	LRLRMVPPPL	PSLNPAPALE	EGQGLTLAAS	CTAEGSPAPS	180
VTWDTEVKGT	TSSRSFKHSR	SAAVTSEFHL	VPSRSMNGQP	LTCVVSHPLG	LQDQRITHIL	240
HVSFLAEASV	RGLEDQNLWH	IGREGAMLC	LSEGQPPPSY	NWTRLDGPLP	SGVRVDGDTL	300
GFPLTTEHS	GIYVCHVSNE	FSRDSQVTV	DVLDPQEDSG	KQVDLVASV	VVVGVIALL	360
FCLLVVVVVL	MSRYHRRKAQ	QMTQKYEEEL	TLTRENSIRR	LHSHHTDPRS	QPEESVGLRA	420
EGHPDSLKDN	SSCSVMSEEP	EGRSYSTLTT	VREIETQTEL	LSPGSGRAEE	EEDQDEGIKQ	480
AMNHVFQENG	TLRAKPTGNG	IYINGRHLV				

# mouse BCX5 Protein sequence (SEQ ID NO:7)

Gene name: mouse\_LNIR; Unigene number: n/a; Probeset Accession #: BF168327; Protein Accession #: n/a; Predicted Signal sequence: 1-27; Predicted TM domains: 346-362; PFAM domains: IgSF\_domains:44-126,166-221,259-313; Summary: This is the mouse orthologue of human BCX5; it is a type Ia TM protein of unknown function.

MPLSLGAEMW	GPEAWLRLLF	LASFTGQYSA	GELETSDEV	VVLGQDAKL	CFYRGDPDEQ	60
VGQVAVARVD	PNEKYPGAGL	LHSKYGLHVN	PAYEDRVEQX	XHETFRSVL	LRNAVQADEG	120

EYECRVSTFP	SGSFQARMRL	RVLVPPPLPSL	NPGPPLEEGQ	ADVAASCTAE	GSPAPSVTWD	180
TEVKGTQSSR	SFTHPRSAAV	TSEFHLVPSR	SMNGQPLTCV	VSHPGLLQDR	RITHTLQVAF	240
LAEASVRGLE	DQNLWQVGRE	GATLKCLSEG	QPPPKYNWTR	LDGPLPSGVR	VKGDTLGFPP	300
LTTEHSGVYX	CHVSNELSSR	DSQVTVEILD	PEDPGKQVDL	VSASVIVIGV	IAALLFCLLV	360
VVVVLMRSYH	RRAQMQMTQK	YEBELTLTRE	NSIRRLHSHH	SDPRSQPEES	VGLRAEGHPD	420
SLKDNSSCSV	MSEPEGRSY	STLTTVREIE	TQTELLSPGS	GRTEEDDDQD	EGIKQAMNHL	480
CRKMGP						

BCZ6 DNA sequence (SEQ ID NO:8)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number: Hs.82065; Probeset Accession #: M57230 / AA406546; Nucleic Acid Accession #: NM\_002184; Coding sequence: 256-3012 (start and stop codons underlined)

GAGCAGCCAA	AAGGCCCGCG	GAGTCGCGCT	GGGCCGCCCC	GGCGCAGCTG	AACCGGGGGC	60
CGCGCCTGCC	AGGCCGACGG	GTCTGGCCCA	GCCTGGCGCC	AAGGGGTTTC	TGCGCTGTGG	120
AGACGCGGAG	GGTCGAGGCG	GCGCGGCCTG	AGTGAAACCC	AATGGAAAAA	GCATGACATT	180
TAGAAGTAGA	AGACTTAGCT	TCAAATCCCT	ACTCCTTCAC	TTACTAATTT	TGTGATTTGG	240
AAATATCCGC	GCAAGATGTT	GACGTTGCAG	ACTTGGGTAG	TGCAAGCCTT	GTTTATTTTC	300
CTCACCATTG	AATCTACAGG	TGAACCTCTA	GATCCATGTG	GTTATATCAG	TCCTGAATCT	360
CCAGTTGTAC	AACCTTCATT	TAATTTCACT	GCAGTTTGTG	TGCTAAAGGA	AAAATGTATG	420
GATTATTTTC	ATGTAAATGC	TAATTACATT	GTCTGGAAAA	CAAACCATT	TACTATTTCCT	480
AAGGAGCAAT	ATACTATCAT	AAACAGAACA	GCATCCAGTG	TCACCTTTAC	AGATATAGCT	540
TCATTAAATA	TTTCAGCTCAC	TTGCAACATT	CTTACATTTC	GACAGCTTGA	ACAGAATGTT	600
TATGGAATCA	CAATAATTTT	AGGCTTGCCT	CCAGAAAAAC	CTAAAAATTT	GAGTTGCATT	660
GTGAACGAGG	GGGAAGAAAT	GAGGTGTGAG	TGGGATGGTG	GAAGGGAAAC	ACACTTGGAG	720
ACAAACTTCA	CTTTAAAAAT	TGAATGGGCA	ACACACAAGT	TTGCTGATTG	CAAAGCAAAA	780
CGTGACACCC	CCACCTCATG	CACTGTTGAT	TATTCTACTG	TGTATTTTGT	CAACATTGAA	840
GTCTGGGTAG	AAGCAGAGAA	TGCCCTTGGG	AAGGTTACAT	CAGATCATAT	CAATTTTGAT	900
CCTGTATATA	AAGTGAAGCC	CAATCCGCCA	CATAATTTAT	CAGTGATCAA	CTCAGAGGAA	960
CTGCTAGTA	TCTTAAAAAT	GACATGGACC	AACCCAAGTA	TTAAGAGTGT	TATAATACTA	1020
AAATATAACA	TTCAATATAG	GACCAAAGAT	GCCTCAACTT	GGAGCCAGAT	TCCTCCTGAA	1080
GACACAGCAT	CCACCCGATC	TTCAATCACT	GTCCAAGACC	TTAAACCTTT	TACAGAATAT	1140
GTGTTTAGGA	TTGCTGTGAT	GAAGGAAGAT	GGTAAGGGAT	ACTGGAGTGA	CTGGAGTGAA	1200
GAAGCAAGTG	GGATCACCTA	TGAAGATAGA	CCATCTAAAG	CACCAAGTTT	CTGGTATAAA	1260
ATAGATCCAT	CCCATACTCA	AGGCTACAGA	ACTGTACAAC	TCGTGTGGAA	GACATTGCCT	1320
CCTTTTGAAG	CCAATGGAAA	AATCTTGGAT	TATGAAGTGA	CTCTCACAAG	ATGGAAATCA	1380
CATTACAAA	ATTACACAGT	TAATGCCACA	AAACTGACAG	TAAATCTCAC	AAATGATCGC	1440
TATCTAGCAA	CCCTAACAGT	AAGAAATCTT	GTTGGCAAAT	CAGATGCAGC	TGTTTAACT	1500
ATCCCTGCCT	GTGACTTTCA	AGCTACTCAC	CCTGTAATGG	ATCTTAAAGC	ATTCCCCAAA	1560
GATAACATGC	TTTGGGTGGA	ATGGACTACT	CCAAGGGGAT	CTGTAAAGAA	ATATATACTT	1620
GAGTGGTGTG	TGTTATCAGA	TAAAGCACCC	TGTATCACAG	ACTGGCAACA	AGAAGATGGT	1680
ACCGTGCAATC	GCACCTATTT	AAGAGGGAAC	TTAGCAGAGA	GCAATATGCTA	TTTGATAACA	1740
GTTACTCCAG	TATATGCTGA	TGGACGAGCA	AGCCCTGAAT	CCATAAAGGC	ATACCTTAAA	1800
CAAGCTCCAC	CTTCCAAAGG	ACCTACTGTT	CGGACAAAAA	AAGTAGGGAA	AAACGAAGCT	1860
GTCTTAGAGT	GGGACCAACT	TCCTGTTGAT	GTTCAGAATG	GATTATACAG	AAATTATACT	1920
ATATTTTATA	GAACCATCAT	TGGAAATGAA	ACTGCTGTGA	ATGTGGATT	TTCCACACA	1980
GAATATACAT	TGCTCTCTTT	GACTAGTGAC	ACATTGTACA	TGGTACGAAT	GGCAGCATA	2040
ACAGATGAAG	GTGGGAAGGA	TGGTCCAGAA	TTCACTTTTA	CTACCCCAAA	GTTTGCTCAA	2100
GGAGAAATTG	AAGCCATAGT	CGTGCTGTT	TGCTTAGCAT	TCCTATTGAC	AACTCTTCTG	2160
GGAGTGCTGT	TCTGCTTTAA	TAAGCGAGAC	CTAATTAAAA	AACACATCTG	GCCTAATGTT	2220
CCAGATCCCT	CAAAGAGTCA	TATTGCCAG	TGGTCACTTC	ACACTCCTCC	AAGGCACAAT	2280
TTTAATTCAA	AAGATCAAAT	GTATTTCAGAT	GGCAATTTCA	CTGATGTAAG	TGTTGTGGAA	2340
ATAGAAGCAA	ATGACAAAAA	GCCTTTTCCA	GAAGATCTGA	AATCATTGGA	CCTGTTCAAA	2400
AAGGAAAAAA	TTAATATCTA	AGGACACAGC	AGTGGTATTG	GGGGGTCTTC	ATGCATGTCA	2460
TCTTCTAGGC	CAAGCATTTC	TAGCAGTGAT	GAAATGAAT	CTTCACAAAA	CACTTCGAGC	2520
ACTGTCCAGT	ATTCTACCGT	GGTACACAGT	GGCTACAGAC	ACCAAGTTCC	GTCAGTCCAA	2580
GTCTTCTCAA	GATCCGAGTC	TACCCAGCCC	TTGTTAGATT	CAGAGGAGCG	GCCAGAAGAT	2640
CTACAATTAG	TAGATCATGT	AGATGGCGGT	GATGGTATTT	TGCCCAGGCA	ACAGTACTTC	2700
AAACAGAACT	GCAGTCAGCA	TGAATCCAGT	CCAGATATTT	CACATTTTGA	AAGGTCAAAG	2760
CAAGTTTCAT	CAGTCAATGA	GGAAGATTTT	GTTAGACTTA	AACAGCAGAT	TTCAGATCAT	2820
ATTTCACAAT	CCTGTGGATC	TGGGCAATG	AAAATGTTTC	AGGAAGTTTC	TGCAGCAGAT	2880
GCTTTTGGTC	CAGTACTGA	GGGACAAGTA	GAAAGATTTG	AAACAGTTGG	CATGGAGGCT	2940
GCGACTGATG	AAGGCATGCC	TAAAGTTTAC	TTACCACAGA	CTGTACGGCA	AGGCGGCTAC	3000
ATGCCCTCAGT	GAAGGACTAG	TAGTTCTCTG	TACAACTTCA	GCAGTACCTA	TAAAGTAAAG	3060
CTAAATGAT	TTTATCTGTG	AATTC				

BCZ6 Protein sequence (SEQ ID NO:9)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number: Hs.82065; Probeset Accession #: M57230 / AA406546; Protein Accession #: NP\_002175; Predicted Signal sequence: 1-22; Predicted TM domains: 625-641; PFAM domains: fibronectin\_type\_III domains: 222-311, 424-509, 519-606; Summary: A type I TM protein; it homodimerizes or heterodimerized to make a functional receptor for IL-6, oncostatin-M, IL-11, LIF, and CNTF.

MLTLQTVVQ	ALFIFLTES	TGELLDPGCV	ISPESPVVQL	HSNFTAVCVL	KEKCMDYFHV	60
NANYIVWKTN	HFTIPKEQYT	IINRTASSVT	FTDIASLNIQ	LTCNLTFTGQ	LEQNVYGITI	120

ISGLPPEKPK	NLSCIVNEGK	KMRCEWDGGR	ETHLETNFTL	KSEWATHKFA	CDKAKRDTPT	180
SCTVDYSTVY	FVNIEVWVEA	ENALGKVTSD	HINFDPVYKV	KPNPPHNLSV	INSEELSSIL	240
KLTWTNPSIK	SVIILKYNIQ	YRTKDASTWS	QIPPEDTAST	RSSFTVQDLK	PFTEYVFRIR	300
CMKEDGKGYW	SDWSEZASGI	TYEDRPSKAP	SFWYKIDPSH	TQGYRTVQLV	WKTLPFPFEAN	360
GKILDYEVTL	TRWKSHLQNY	TVNATKLTVN	LTNDRYLATL	TVRNLVGKSD	AAVLTIACD	420
FQATHPVMDL	KAFPKDNMLW	VEWTTPRESV	KKYILEWCVL	SDKAPCITDW	QQEDGTVHRT	480
YLRGNLAESK	CYLITVTPVY	ADGPGSPESI	KAYLKQAPPS	KGPTVRTKKV	GKNEAVLEWD	540
QLPVDVQNGF	IRNYTIFYRT	IIGNETAVNV	DSSHTEYTLS	SLTSDTLYMV	RMAAYTDEGG	600
KDGPEFTFTT	PKFAQGEIEA	IVVPVCLAFI	LTLLGLVLF	FNKRDLIKHH	IWPNVPDPSK	660
SHIAQWSPHT	PPRHNFNISK	QMYSDGNFTD	VSVVEIEAND	KKPFPEDLKS	LDLFKKEKIN	720
TEGHSSGIGG	SSCMSSSRPS	ISSSDENESS	QNTSSTVQYS	TVVHSGYRHO	VPSVQVFSRS	780
ESTQPLLDSE	ERPEDLQLVD	HVDGGDGILP	RQYFKQNC	QHSSPDISH	FERSKQVSSV	840
NEEDFVRLKQ	QISDHISQSC	GSGQMKMFQE	VSAADAFGPG	TEGQVERFET	VGMEAATDEG	900
MPKSYLPQTV	RQGGYMPQ					

BFG4 DNA sequence (SEQ ID NO:10)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;  
Nucleic Acid Accession #: AB020689; Coding sequence: 108-2777 (start and stop codons  
underlined)

GAACCTATGT	AGCCTCATTA	TCCCGCTCCG	TGAGGTGACA	ATTGTGGAAG	AGGCAGACAG	60
CTCCAGTGTG	CTCCCCAGTC	CCTTATCACA	TCAGCACCCG	AAACAGGATG	ACCTTCTTAT	120
TTGCTCAACT	GAAAGATAGA	GACTTTCTAG	TGCAGAGGAT	CTCAGATTTC	CTGCAACAGA	180
CTACTTCCAA	AATATATTCT	GACAAGGAGT	TTGCAGGAAG	TTACAACAGT	TCAGATGATG	240
AGGTGTACTC	TCGACCCAGC	AGCCTCGTCT	CCTCCAGCCC	CCAGAGAAGC	ACGAGCTCTG	300
ATGCTGATGG	AGAGCGCCAG	TTTAACCTAA	ATGGCAACAG	CGTCCCCACA	GCCACACAGA	360
CCCTGATGAC	CATGTATCGG	CGGCGGTCTC	CCGAGGAGTT	CAACCCGAAA	TTGGCCAAAG	420
AGTTTCTGAA	AGAGCAAGCC	TGGAAGATTC	ACTTTGCTGA	GTATGGGCAA	GGGATCTGCA	480
TGTACCGCAC	AGAGAAAACG	CGGGAGCTGG	TGTTGAAGGG	CATCCCGGAG	AGCATGCGTG	540
GGGAGCTCTG	GCTGCTGCTG	TCAGGTGCCA	TCAATGAGAA	GGCCACACAT	CCTGGGTACT	600
ATGAAGACCT	AGTGGAGAAG	TCCATGGGGA	AGTATAATCT	CGCCACGGAG	GAGATTGAGA	660
GGGATTTTACA	CCGCTCCCTT	CCAGAACACC	CAGCTTTTCA	GAATGAAATG	GGCATTGCTG	720
CACATAAGGAG	AGTCTTAACA	GCTTATGCTT	TTCGAAATCC	CAACATAGGG	TATTGCCAGG	780
CCATGAATAT	TGTCACCTTCA	GTGCTGCTGC	TTTATGCCAA	AGAGGAGGAA	GCTTTCTGGC	840
TGCTTGTGGC	TTTGTGTGAG	CGCATGCTCC	CAGATTACTA	CAACACCAGA	GTTGTGGGTG	900
CACCTGGTGA	CCAAGGTGTC	TTTGAGGAGC	TAGCACGAGA	CTACGTCCCA	CAGCTGTACG	960
ACTGCATGCA	AGACCTGGGC	GTGATTTCCA	CCATCTCCCT	GTCTTGGTTC	CTCACACTAT	1020
TTCTCAGTGT	GATGCCCTTT	GAGAGTGCGA	TTGTGGTTGT	TGACTGTTTC	TTCTATGAAG	1080
GAATTAAGT	GATATTCCAG	TTGGCCCTAG	CTGTGCTGGA	TGCAATGTG	GACAACTGT	1140
TGAACGTCAA	GGATGATGGG	GAGGCCATGA	CCGTTTGGG	AAGGTATTTA	GACAGTGTGA	1200
CCAATAAAGA	CAGCACACTG	CCTCCCATTC	CTCACCTCCA	CTCCTTGCTC	AGCGATGATG	1260
TGGAACCTTA	CCCTGAGGTA	GACATCTTTA	GACTCATCAG	AACTTCTTAC	GAGAAATTCG	1320
GAACATATCC	GGCAGATTTG	ATTGAACAGA	TGAGATTCAA	ACAGAGACTG	AAAGTGATCC	1380
AGACGCTGGA	GGATACTACG	AAACGCAACG	TGGTACGAAC	CATTGTGACA	GAAACTTCCT	1440
TTACCATTTA	TGAGCTGGAA	GAACCTTATG	CTCTTTTCAA	GGCAGAACAT	CTCACCAGCT	1500
GCTACTGGGG	CGGGAGCAGC	AACGCGCTGG	ACCGGCATGA	CCCCAGCCTG	CCCTACCTGG	1560
AACAGTATCG	CATTGACTTC	GAGCAGTTCA	AGGGAATGTT	TGCTCTTCTC	TTTCTTGGG	1620
CATGTGGAAC	TCACCTGAC	GTTCTGGCCT	CCCGCTTGT	CCAGTTATTA	GATGAAATG	1680
GAGACTCTTT	GATTAACTTC	CGGGAGTTTG	TCTCTGGGCT	AAGTCTGCA	TGCCATGGGG	1740
ACCTCACAGA	GAAGCTCAA	CTCCTGTACA	AAATGCACGT	CTTGCCTGAG	CCATCCTCTG	1800
ATCAAGATGA	ACCAGATTCT	GCTTTTGAAG	CAACTCAGTA	CTTCTTTGAA	GATATTACCC	1860
CAGAAATGTAC	ACATGTTGTT	GGATTGGATA	GCAGAAAGCA	ACAGGTGCA	GATGATGGCT	1920
TTGTTACGGT	GAGCCTAAAG	CCAGACAAAG	GGAAGAGAGC	AAATCCCAA	GAAATTCGTA	1980
ATTATTGTAG	ACTGTGGACT	CCAGAAATA	AATCTAAGTC	AAAGAATGCA	AAGGATTATC	2040
CCAAATTAAT	TCAGGGGCG	TTCATTGAAC	TGTGTAAGAC	AATGTATAAC	ATGTTACGCG	2100
AAGACCCCAA	TGAGCAGGAG	CTGTACCATG	CCACGGCAGC	AGTGACCAGC	CTCCTGCTGG	2160
AGATTGGGGA	GGTCGGCAAG	TTGTTGCTGG	CCAGCCTGC	AAAGGAGGGC	GGGAGCGGAG	2220
GCAGTGGGCC	GTCCTGCGAC	CAGGGCATCC	CAGGCGTGCT	CTTCCCAAG	AAAGGGCCAG	2280
GCCAGCCTTA	CGTGGTGGAG	TCTGTTGAGC	CCCTGCCGGC	CAGCCTGGCC	CCCACAGCG	2340
AGGAACACTC	CCTTGGAGGA	CAAATGGAGG	ACATCAAGCT	GGAGGACTCC	TCGCCCCGGG	2400
ACAACGGGGC	CTGCTCTCTC	ATGCTGATCT	CTGACGACGA	CACCAAGGAC	GACAGCTCCA	2460
TGCTCTCATA	CTCGGTGCTG	AGTGCCGGCT	CCCACGAGGA	GGACAAGCTG	CACTGCGAGG	2520
AAATCGGAGA	GGACACGGTC	CTGGTGCGGA	GCGGCCAGGG	CACGGCGGCA	CTGCCCCGGA	2580
GCACACGCTT	GGACCGGGAG	TGGGCCATCA	CCTTCGAGCA	GTTCTGGGCC	TCCCTCTTAA	2640
CTGAGCCTGC	CCTGGTCAAG	TACTTTGACA	AGCCCGTGTG	CATGATGGCC	AGGATTACCA	2700
GTGCAAAAAA	CATCCGGATG	ATGGGCAAGC	CCCTCACCTC	GGCCAGTGAC	TATGAAATCT	2760
CGGCCATGTC	CGGCTGACAC	GGGCGCCTTC	CCGGGGGAGT	GGGAGGAGAG	GGAGGGGAGG	2820
GATTTTAT	GTTCTCTCTG	GTTGAGTTTT	TTCTTTCTTT	CTTTTAAATT	AAATATTTAT	2880
TAGTACCTGG	AATTGAAGCC	TAGTGTTTTC	ATAATGTAAT	TCAATGAAAA	CTGTTGGAGA	2940
GAATATTAAT	CACCTCAATG	TAGGTACATT	ACACTCTTGT	TGCGGGGAGG	GGATTTACCA	3000
GAATACAGTT	TATTTCTGTG	ATTCTAAAAA	ACAAAAAGAT	GAATCTGTCA	GTGATATGTG	3060
TGTATTATAA	CTTATTAATC	TTGCTGTTGA	GCTGTATACA	TGGTTTAAAA	AATAGTACTG	3120
TTTAATGCTA	AGTAAGGCGA	CAGTCATTTG	TGTATTCAGG	CTTTTAAAT	AAAATTAGAG	3180
CTGTAAGGAA	AATGAAAAAG	CACAAATGCA	AGACTGTTCT	TAAATGGAAG	GCATAGTCAG	3240
CGAGGTAATA	TCCTATACCA	CTTTAGGAAG	TATTAATAAT	ATTTTAAAGA	TTTGAAATAT	3300
ATTTTCATAGA	AGTCTCTAT	TCAAAATCAT	ATTCCACAGA	TGTTCCCTTT	CAAAGGAAAA	3360



ACATTGTTGGG	TTCTAAACAG	TTATGAAAGT	AAGTGATTTT	TACATGATTG	CAGAATAACA	3420
CTTGTATTGA	CCAATTTAGA	CAGATACCCAG	ACCAATTTTG	CATTAAAGAA	ATTGTTCTGA	3480
TTATTTCAGT	CAACTCATT	GAATTCAGTG	AAAAGTAACA	GTCTTTTGTC	ACAGAGAATC	3540
TGAAAGTAGC	AGCAAAGACA	GAGGGCTCAT	GACAGGTTTT	TGCTTTTGCT	TTGCTTTTGT	3600
TTTTGAAAGA	GTAAGAATAC	TGATGCTTCT	GATACTGGAT	GTTTAGCTTC	TTACTGCAAA	3660
AACATAAGTA	AAACAGTCAA	CTTTACCATT	TCCGTATTCT	CCATAGATTG	AAGAAATTTA	3720
TACCACATAT	CGCATATGAC	CATCTTTCCA	TCAAATCAAT	GTAGAGATAA	TGTAAACTGA	3780
AAAAAAATCT	GCAAGATAAT	GTAAGTGAAT	GTTTTAAAAA	CAGAACTTGT	CACTTTATAT	3840
AAAAAGAATG	TATGCTCTAT	TTCTGGAATG	GATGTGGAAT	TGAAAGCTAG	CGCACCTGCA	3900
CTTTGAATTC	TTGCTTCTTT	TTTATTACTG	TTATGATTTT	GCTTTTGTCA	GATGTTGGAC	3960
GATTTTCTCT	TCTGATTGTT	GAATTCATAA	TCATGGTCTC	ATTTCTTTTG	CTTCTTTGGA	4020
ATATTCTTTT	CAACACATTG	CTTTATTTTA	TTATACATTG	TGTCCTTTT	TTAGCTATTG	4080
CTGCTGTTGT	TTTTTATTCT	ATTTACAGGA	TGATTTTAA	ACTGTCAAAT	GAAGTAGTGT	4140
TAACCTCAAA	TAGGCTAAAT	GTGAACAAAT	AAAATACAGC	AAATACTCAG	AAAAAATAAA	4200
AAAAAATAAA	AAAAA					

BFG4 Protein sequence (SEQ ID NO:11)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;  
 Protein Accession #: BAA74905; Signal sequence: none; Predicted TM domains: 302-318; PFAM  
 domains: TBC\_domain: 135-347; Summary: a Type II membrane protein, likely localized to the  
 peroxisome.

MTFLFANLKD	RDFLVQRISD	FLQQTTSKIY	SDKEFAGSYN	SSDDEVYSRP	SSLVSSSPQR	60
STSSDADGER	QFNLMGNSVP	TATQTLMTMY	RRRSPEEFNP	KLAKEFLKEQ	AWKIHFAEYV	120
QGICMYRTEK	TRELVLKGIP	ESMRGELWLL	LSGAINNEKAT	HPGYEDLVE	KSMGKYNLAT	180
EEIERDLHRS	LPEHLPQNE	MGIAALRRVL	TAYAFRNPNI	GYCQAMNIVT	SVLLLYAKEE	240
EAFWLLVALC	ERMLPDYNT	RVVGALVDQ	VFEELARDYV	PQLYDCMQDL	GVISTISLSW	300
FLTLFLSVMP	FESAVVVVDC	FFYEGIKVIF	QLALAVLDAN	VDKLLNCKDD	GEAMTVLGRY	360
LDSVTNKDST	LPPPHLHSL	LSDDVEPYPE	VDIFRLIRTS	YEKFGTIRAD	LIEQMRFKQR	420
LKVIQTLEDT	TKRNVVRTIV	TETSFTIDEL	EELYALFKAE	HLTSCYWGGG	SNALDRHDP	480
LPYLEQYRID	FEQFKGMFAL	LFPWACGTHS	DVLASRLFQL	LDENGDSLIN	FREFVSGLSA	540
ACHGDLTEKL	KLLYKMHVLP	EPSSDQDEPD	SAFEATQYFF	EDITPECTHV	VGLDSRSKQG	600
ADDGFTVSL	KPDGKGRANS	QENRNYLRWL	TPENKSKSKN	AKDLPKLNQ	QFIELCKTMY	660
NMFSEDPNEQ	ELYHATAAVT	SLLEIGEVEG	KLFVAQPAKE	GGSGSGSPSC	HQGIPGVLP	720
KKGPGQPVV	ESVEPLPASL	APDSEHSLG	GQMEDIKLED	SSPRDNGACS	SMLISDDDTK	780
DDSSMSSYSV	LSAGSHEEDK	LHCEIGEDT	VLVRSGQGT	ALPRSTSLDR	DWAITFEQFL	840
ASLLTEPALV	KYFDKPVCM	ARITSAKNIR	MMGKPLTSAS	DYEISAMSG		

BCU7 DNA sequence (SEQ ID NO:12)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Nucleic Acid  
 Accession #: n/a; Coding sequence: 1-573 (stop codon underlined)

TATTTTATTT	TCCAGGCTAA	AGCAAATGAA	AGTTTGCTGG	TATCAACACA	GCCTGCCATA	60
TTTTTCACAG	CATGCAACAA	TGGTGCTAGG	ATAGCTATTT	CTTACTGTAA	TTGCCAGAGG	120
CAGAAATGGT	CTGGGTATAA	GCTATTTTCT	AAAAGCAGCT	TTAAATTGTC	AGTATTAAGG	180
TTTTCATGTG	GAAAGGTGTC	ATTCAAAAAA	AAAGTAATTG	GCATACATAT	TCCACATCAT	240
CGATCCTCTC	TGTGGTGTTA	ATTTTCTTAT	ATGACCAGTA	GAAAAATTTT	AATATTCTCA	300
CAATATAGGT	TTTGGGGCTT	CCATATCATC	AAAAGACTGA	AAAATTATAA	TTTTAGAATT	360
AAACTGATGG	ATTTTATTAT	AGAATTATCT	GTGAGTTGTG	TAGACACAGT	CTTAATGTTT	420
TTATTGCTTC	TGTTTCTGTT	TCTAAGTATC	ATCCCTTCTG	TGGCCCATCA	CGCAGCAGAG	480
TTGCCCTACA	AATTTTCTTT	GGCAGCGCCA	<u>TAA</u> CATTCAT	TTAAAAAGTT	TATGAAAACA	600
TTCATTTGAA	AGTTTCCATG	AGCTTTAGCA	CAGAGTTGAC	CAAACACTGG	CGTAAGTTCA	660
ATTTACACAG	AATATTTGAA	TTGAAACAAT	AGAAATTTT	CTCATAATAT	ATACCTATGT	720
GAAACCAACT	TATCTGCATA	ATTAAATCTA	ATACATATTT	AAGCCAGTTT	AAGTGCTTTG	780
TGTTGATGCC	ATGCTTATCA	AATACATGCA	CAAGCTAAAC	ATAATTTGAA	TGGGCTATG	840
AAGGAAAAAT	AATGCTTAGA	CTTTGGTGTA	GGTTCTTCT	GTGTAGCCAT	ATACCCAGGC	900
TCTGCAGTAT	CGAAGGATGC	AAATGTTGAC	ATAGATGGAA	GCTCTTACCT	ACCAAAGTGT	960
TTAGGAAGGA	TAAAGTTACA	TTTGTCTTAA	TTTCTAACAT	TATCTTTGCT	TTTATGTTTC	1020
ATAAAAAATTT	GTCATTATTT	ATGCTGGTGA	AACGTATAAT	CACATCCAAT	TATTTGAACA	1080
CATGCAAAAT	AATTTTCTTA	ATTATGTTAT	TGTTTAAATT	TGACTTATGG	GAGATCAGTC	1140
AAAAACTTAG	AAGGTTTAAC	ACTTCACTGA	TTAATGGTGC	TGAAAACACG	TTACAATTAC	1200
CACATATCCT	TGCTATAAGT	TTTGAAGTTT	CTTAGCAATT	AAAGTTTTTT	TATTCAAGTGT	1260
GAAGTGTGAG	TATCTATTCT	GGTGCTAAAT	GTATGGTGCT	AAATGAATTG	TTAGTGTGTA	1320
TGGCTTTAGT	AATGCTCCTT	TTATTCATTG	CTAAATTTAG	TGTTATCCAT	TTGATTCTCT	1380
ATTACAGAAAT	ATCAATAAAA	TCCTATGTTA	AATTAATCTT	TACCAAAAAAC	AGGCAAGTTA	1440
ACTCTGTTGT	TTTAATTCAA	CAGTCCAACA	TTATTTAGGT	GTTACAGAGT	GTAATATAT	1500
TTCTTTGGGA	GTTATTTTCT	TTTTTAAATC	TTTTTATAGC	TTGGCAATGT	CCAAAGTCAA	1560
ATATCACCTA	AACTGGTTAG	ATTACTTCTA	CAGCTAATAA	TATTGCAAGC	ACTGGCGCCC	1620
TCTGGTGGTT	ATGAAGACAA	ATTCTTAATG	GCTACTTGAC	CTACAGCAAA	AGCCATTTCT	1680
GTACCATAAA	AAATTTGTTG	GCAATATTAG	AATTATCATA	TGTTTCTTAC	ATCTGACAGC	1740
ACCTAAAAATG	TTTGATAATA	TTAACATGTA	TCTAAGAGGA	AAAAAGAGTT	AATATATTCT	1800
GGCACCCACT	TTCCAGTAGT	TGTTTCCAT	GATTTTCCAG	TTCTGAGGCA	CTTATTAAG	1860
TGCTTTTTTT	TTTCTGAATT	AATTAGGTAT	TGGTAAAAATA	TATTTTTAAA	TTAGTTAGC	1920
TTTATAAACA	CAATTAGAAAT	TACAATTAAT	TAACAGAGGT	ATAATGTCT	CACTTTCAGA	1980
AGTGATCATT	TATTTTATT	TAGCACAGGT	CATAAGAAAA	ATATATAGAA	AAATAATCAA	2040

TTTCATATAT	AAAAGGATTA	TTTCTCCACC	TTTAATTATT	GGCCTATCAT	TTGTTAGTGT	2100
TATTTGGTCA	TATTATTGAA	CTAATGTATT	ATTCCATTCA	AAGTCTTTCT	AGATTTAAAA	2160
ATGTATGCAA	AAGCTTAGGA	TTATATCATG	TGTAACATAT	ATAGATAACA	TCCTAAACCT	2220
TCAGTTTAGA	TATATAATTG	ACTGGGTGTA	ATCTCTTTTG	TAATCTGTTT	TGACAGATTT	2280
CTTAAATTAT	GTTAGCATAA	TCAAGGAAGA	TTTACCTTGA	AGCACTTTCC	AAATTGATAC	2340
TTTCAAACCT	ATTTTAAAGC	AGTAGAACCT	TTTCTATGAA	CTAAATCACA	TGCAAAACTC	2400
CAACCTGTAG	TATACATAAA	ATGGACTTAC	TTATTCCTCT	CACCTTCTCC	AGTGCCCTAGG	2460
AATATTCCTC	TCTGAGCCCT	AGGATTGATT	CTATCACACA	GAGCAACATT	AATCTAAATG	2520
GTTTAGCTCC	CTCTTTTTTC	TCTAAAAACA	ATCAGCTAAT	AAAAAAAAAA	TTTGAGGGCC	2580
TAAATTATTT	CAATGGTTGT	TTGAAATATT	CAGTTCAGTT	TGTACCTGTT	AGCAGTCTTT	2640
CAGTTTGGGG	GAGAATTAAA	TACTGTGCTA	AGCTGGTGCT	TGGATACATA	TTACAGCATC	2700
TTGTGTTTTA	TTTGACAAAC	AGAATTTTGG	TGCCATAATA	TTTTGAGAAT	TAGAGAAGAT	2760
TGTGATGCAT	ATATATAAAC	ACTATTTTTA	AAAAATATCT	AAATATGTCT	CACATATTTA	2820
TATAATCCTC	AAATATACTG	TACCATTTTA	GATATTTTTT	AAACAGATTA	ATTTGGAGAA	2880
GTTTTATTCA	TTACCTAATT	CTGTGGCAAA	AATGGTGCCT	CTGATGTTGT	GATATAGTAT	2940
TGTCAGTGTG	TACATATATA	AAACCTGTGT	AAACCTCTGT	CCTTATGAAC	CATAACAAAT	3000
GTAGCTTTTT	AAAGTCCATT	GTATTGTTTT	TTCTTTCAAT	AAAAGAGTAT	AATTAATTGG	3060
TTGTTTTTGA						

BCU7 Protein sequence (SEQ ID NO:13)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 125-141, 154-170; PFAM domains: none; Summary: A type III membrane protein, highly overexpressed in breast cancer and prostate cancer; unknown function.

YFIFQAKANE	SLLVSTQPAI	FFTACNNGAR	IAISYCNCQR	QKWSGYKLPH	KSSFKLSVLR	60
FSCGKVSFFK	KVIGIHIPHH	RSSLWCXFFY	MTSRKILIFS	QYRFWGFHII	KRLKNYNFRI	120
KLMDFIIELS	VSCVDTVLMF	LVMTDKFAQK	MWMKPLLLLL	LLLLFSLSLI	IPSVAHAAAE	180
LPYKFHLAAP						

BFA1 DNA sequence (SEQ ID NO:14)

Gene name: calsyntenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Nucleic Acid Accession #: NM\_022131; Coding sequence: 11-2878 (start and stop codons underlined)

TGCTGCGAGG	<u>ATGCTGCCTG</u>	GGCGGCTGTG	CTGGGTGCCG	CTCCTGCTGG	CGCTGGGCGT	60
GGGGAGCGGC	AGCGGCGGTG	CGGGGACAG	CCGGCAGCGC	CGCCTCCTCG	CGGCTAAAGT	120
CAATAAGCAC	AAGCCATGGA	TCGAGACTTC	ATATCATGGA	GTCATAACTG	AGAACAATGA	180
CACAGTCATT	TTGGACCCAC	CACCTGGTAGC	CCTGGATAAA	GATGCACCGG	TTCTTTTTCG	240
AGGGGAAATC	TGTGCGTTCA	AGATCCATGG	CCAGGAGCTG	CCCTTTGAGG	CTGTGGTGCT	300
CAACAAGACA	TCAGGAGAGG	GCCGGCTCCG	TGCCAAGAGC	CCCATTTGACT	GTGAGTTGCA	360
GAAGGAGTAC	ACATTTCATCA	TCCAGGCCTA	TGACTGTGGT	GCTGGGCCCC	ACGAGACAGC	420
CTGGAAAAAG	TCACACAAGG	CCGTGGTCCA	TATACAGGTG	AAGGATGTCA	ACGAGTTTGC	480
TCCCACCTTC	AAAGAGCCAG	CTCAAGGCG	TGTTGTGACG	GAGGGCAAGA	TCTATGACAG	540
CATTCTGCAG	GTGGAGGCCA	TTGACGAGGA	CTGCTCCCCA	CAGTACAGCC	AGATCTGCAA	600
CTATGAAATC	GTCAACACAG	ATGTGCCTTT	TGCCATCGAC	AGAAATGGCA	ACATCAGGAA	660
CACTGAGAAG	CTGAGCTATG	ACAAACAACA	CCAGTATGAG	ATCCTGGTGA	CCGCCTACGA	720
CTGTGGACAG	AAGCCCGCTG	CTCAGGACAC	CCTGGTGACG	GTGGATGTGA	AGCCAGTTTG	780
CAAGCCTGGC	TGGCAAGACT	GGACCAAGAG	GATTGAGTAC	CAGCCTGGCT	CCGGGAGCAT	840
GCCCTGTGTC	CCCAGCATCC	ACCTGGAGAC	GTGCGATGGA	GCCGTGTCTT	CCCTCCAGAT	900
CGTCACAGAG	CTGCAGACTA	ATTACATTGG	GAAGGGTTGT	GACCGGGAGA	CCTACTCTGA	960
GAAATCCCTT	CAGAAGTTAT	GTGGAGCCTC	CTCTGGCATC	ATTGACCTCT	TGCCATCCCC	1020
TAGCGCTGCC	ACCAACTGGA	CTGCAGGACT	GCTGGTGGAC	AGCAGTGAGA	TGATCTTCAA	1080
GTTTGACGGC	AGGCAGGGTG	CCAAAATCCC	CGATGGGATT	GTGCCCAAGA	ACCTGACCGA	1140
TCAGTTTCAC	GTGATGAAAC	CGGCCCCAGC	CTGGTGTGTA	GAGCCGAGAA		1200
GGAAACCATC	CTCTGCAACT	CAGACAAAAC	CGAAATGAAC	CGGCATCACT	ATGCCCTGTA	1260
TGTGCACAAC	TGCCGCCTCG	TCTTTCTCTT	GCGGAAGGAC	TTCGACCAGG	CTGACACCTT	1320
TCGCCCCGCG	GAGTTCCACT	GGAAGCTGGA	TCAGATTTGT	GACAAAGAGT	GGCACTACTA	1380
TGTCATCAAT	GTGGAGTTTC	CTGTGGTAAC	CTTATACATG	GATGGAGCAA	CATATGAACC	1440
ATACCTGTGT	ACCAACGACT	GGCCCCATTCA	TCCATCTCAC	ATAGCCATGC	AACTCACAGT	1500
CGGCGCTTGT	TGGCAAGGAG	GAGAAGTCAC	CAAAACCACAG	TTTGCTCAGT	TCTTTCATGG	1560
AAGCCTGGCC	AGTCTCCACCA	TCCGCCCTGG	CAAAATGGAA	AGCCAGAAGG	TGATCTCCTG	1620
CCTGCAGGCC	TGCAAGGAAG	GGCTGGACAT	TAATTCCTTG	GAAAGCCTTG	GCCAAGGAAT	1680
AAAGTATCAC	TTCAACCCCT	CGCAGTCCAT	CCTGGTGATG	GAAGGTGACG	ACATTGGGAA	1740
CATTAAACCG	GCTCTCCAGA	AAGTCTCCTA	CATCAACTCC	AGGCAGTTCC	CAACGGCGGG	1800
TGTGCGGCGC	CTCAAAGTAT	CCTCCAAAGT	CCAGTGCTTT	GGGGAAGACG	TATGCATCAG	1860
TATCCCTGAG	GTAGATGCCT	ATGTGATGGT	CCTCCAGGCC	ATCAGAGCCC	GGATCACCCCT	1920
CCGGGGCACA	GACCACTTCT	GGAGACCTGC	TGCCAGTTT	GAAAGTGCCA	GGGGAGTGAC	1980
CCTCTTCCCT	GATATCAAGA	TTGTGAGCAC	CTTCGCCAAA	ACCGAAGCCC	CCGGGGACGT	2040
GAAACCACAC	GACCCCAAAT	CAGAAGTCTT	AGAGGAAATG	CTTCATAACT	TAGATTTCTG	2100
TGACATTTTG	GTGATCGGAG	GGGACTTGGA	CCCAAGGCAG	GAGTGCTTGG	AGCTCAACCA	2160
CAGTGAGTGT	CACCAACGAC	ACCTGGATGC	CACTAATTCT	ACTGCAGGCT	ACTCCATCTA	2220
CGGTGTGGGC	TCCATGAGCC	GCTATGAGCA	GGTGCTACAT	CACATCCGCT	ACCGCAACTG	2280
GCGTCCGGCT	TCCCTTGAGG	CCCGGCGTTT	CCGGATTAAG	TGCTCAGAAC	TCAATGGGCG	2340
CTACACTAGC	AATGAGTTCA	ACTTGGAGGT	CAGCATCCTT	CATGAAGACC	AAGTCTCAGA	2400
TAAGGAGCAT	GTCAATCATC	TGATTGTGCA	GCCTCCCTTC	CTCCAGTCTG	TCCATCATCC	2460
TGAGTCCCGG	AGTAGCATCC	AGCACAGTTC	AGTGGTCCCA	AGCATTGCCA	CAGTGGTCAT	2520

CATCATCTCC	GTGTGCATGC	TTGTGTTTGT	CGTGGCCATG	GGTGTGTACC	GGGTCCGGAT	2589
CGCCCAACCAG	CACCTTCATCC	AGGAGACTGA	GGCTGCCAAG	GAATCTGAGA	TGGACTGGGA	2640
CGATTCTGCG	CTGACTATCA	CAGTCAACCC	CATGGAGAAA	CATGAAGGAC	CAGGGCATGG	2700
GGAAAGATGAG	ACTGAGGGAG	AAGAGGAGGA	AGAAGCCGAG	GAAGAAATGA	GCTCCAGCAG	2760
TGGCTCTGAC	GACAGCGAAG	AGGAGGAGGA	GGAGGAAGGG	ATGGGCAGAG	GCAGACATGG	2820
GCAGAAATGGA	GCCAGGCAAG	CCCAGCTGGA	GTGGGATGAC	TCCACCCTCC	CCTACTAGTG	2880
CCCAGGGGTC	TGCTGCCTGG	CCCACATGTC	CCTTTTGTA	ACCCTGACCC	AGTGTATGCC	2940
CATGTCTATC	ATACCTCACC	TCTGATGTCT	GTGACATGTC	TGGGAAGGCC	TTCTCCAGCT	3000
TCCTGGAGCC	CACCCCTTAA	GCCTTGGGCA	CTCCCTGTGT	TTCATCCATG	GGGAAGTTCC	3060
AAGAAGCCCA	GCATGGCCAT	CAGTGAGGAC	TTCAGGGTAG	ACTTTGTCCT	GTAGCCTCCA	3120
CTTCTGCCCT	AAGTTCCCCA	GCATCCTGAC	TACCTGTCTG	CAGAGTTTGC	CTTTGTTTTT	3180
TCCTGCAGGG	AAGAAGGCC	ACCTTTGTGT	CACCTACCTC	CCCAGGCTCA	GAGTCCCCAA	3240
GGCCCTGGGG	TTCCAACCTCA	CTGTGCGTCT	CCTCCACACA	GACCAGTAGG	TTCTCCTATG	3300
CTGACTCCAG	GTTGCTTTCAT	ACAAGGAGGG	TGGTTGAACT	TCACACACGT	AAGGTCTTAG	3360
TGCTTAACAG	TTTAAAGGAA	AGTCCTTGTT	GAGGCAGAAC	TAAGTTTACA	GGGAAAGGTA	3420
CACACATTCT	CTCTCTCTCT	CTCTCTCTGT	CTATCTAGTT	CCCCAGCTTG	GAGAGCCTTT	3480
CCCCTTGCTT	CTTTCTGAGG	CCATATAAGC	TTATAAGAAA	AGTCCCAAAC	CAAGAATAGG	3540
TCCTTGGGCA	CACGAGGGT	CTGATCCCCC	ATCAGAGCTA	TCTGAGCCTG	CCTGTCTGGG	3600
CACCTGCTGC	AACCATGCAG	CTACCTTGCC	AGGGGCACCT	AGCAAACAGA	ACCACAGGGC	3660
CCAGGAGGCA	TTCCACACAG	GCACTGCCCC	AGGACAACAC	AACAAGGACA	GTCACAACAA	3720
GGACAACAAG	GACACAACAC	AACACACAAC	AAGGACAGTC	ACAACAAGCC	TAGAGCCAGA	3780
AAGCAGATGG	AAAAGCTAAT	GAGGTCAAAC	GTAGGCTTCA	TGGTGGGTGG	AGTGGGGGTG	3840
GCTGGGCTCC	CCCAGGACAG	AGGGGACCTC	GAGGTGGGCA	AGGCTCTCAC	CACCTCAGCT	3900
TATGGTCCCT	TATCTCCTAT	CTTCCCTCTT	GAGAAAATAC	ACGCTTTCTG	CATGTATTAG	3960
AAACGCACGA	GCTCCACCAA	GTCTACAATG	AAAGTTTGAA	ATTTAACTGC	AAGGAATTAG	4020
AAGCATATTG	GCAATCATTT	CAGCTTCTTC	TTTCTTCTGC	TCATAAAAGG	AGGAACACTT	4080
TAGATAGAGG	GCAAATATAT	CTGAAAACCT	AATTTCTTTC	TTTTTTTGAT	AAGGAAATCT	4140
TTTCCATCTC	CATCCTAACA	TGCACAACCT	GTGAAGAGAA	TTGTTTCTAT	AGTAACTGGT	4200
CTGTGATCTT	TTGTGGCCAA	GAGAATAGCA	GGCAAGAATT	AGGGCCTTGA	CAGAATTTC	4260
ACGAAGCTCT	GAGAACATGT	TTGTTTCGAA	TGCTGTATTG	CTCTTTGTCA	TCAATGTGTA	4320
TGCTCTGTCC	CCATCCTTCA	CTCCTCCTCA	AGCTCACACC	AATTGGTTTTG	GCACAGGCAC	4380
AGAGCTGGTC	CCTAGTTAAG	TGGCATTAT	GTTAAAAAAA	A		

# BFA1 Protein sequence (SEQ ID NO:15)

Gene name: calsyntenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Protein Accession #: NP\_071414; Predicted Signal sequence: 1-20; Predicted TM domains: 832-848; PFAM domains: cadherin\_domains: 48-151, 165-254; Summary: A type I membrane protein; a member of the calsyntenin family; is related to the FAT tumor suppressor; is likely an adhesion molecule important in mammalian developmental processes and cell communication.

MLPGRLCWVP	LLLALGVGSG	SGGGGDSRQR	RLLAAKVKNH	KPWIETSYHG	VITENNDTVI	60
LDPPLVALDK	DAPVPFAGEI	CAFKIHGQEL	PFEAVVLNKT	SGEGRLRKAS	PIDCELQKEY	120
TFIIQAYDCG	AGPHETAWKK	SHKAVVHIQV	KDVNEFAPTF	KEPAYKAVVT	EGKIYDSILQ	180
VEAIDEDCSP	QYSQICNYEI	VTTDVPFAID	RNGNIRNTEK	LSYDKQHQYE	ILVTAYDCGQ	240
KPAAQDTLVQ	VDVKPVCKPG	WQDWTKRIEY	QPGSGSMPLE	PSIHLETCDG	AVSSLQIVTE	300
LQTNYIGKGC	DRETYSEKSL	QKLCGASSGI	IDLPLSPSAA	TNWTAGLLVD	SSEMIFKFDG	360
RQGAIPDGI	VPKNLTDQFT	ITMWMKHGPS	PGVRAEKETI	LCNSDKTEMN	RHHYALYVHN	420
CRLVFLLRKD	FDQADTFRPA	EFHWKLDQIC	DKEWHYYVIN	VEFPVVTLYM	DGATYEPYLV	480
TNDWPIHPSH	IAMQLTVGAC	WQGGVETKPG	FAQFFHGLSL	SLTIRPGKME	SQKVISCLOA	540
CKEGLDINSL	ESLGGIKYH	FNPSQSILVM	EGDDIGNINR	ALQKVSYINS	RQFPTAGVRR	600
LKVSSKVQCF	GEDVCISIFE	VDAYVMVLQA	IEPRITLRGT	DHFWRPAAQF	ESARGVTLPF	660
DIKIVSTFAK	TEAPGDVKT	DPKSEVLEEM	LHNLDPCDIL	VIGGDLDPHQ	ECLELNHSEL	720
HQRHLDATNS	TAGSYIYGVH	SMSRYEQVLH	HIRYRNWRPA	SLEARRFRIK	CSELNGRYTS	780
NEFNLEVSIL	HEDQVSDKEH	VNHLIVQPPF	LQSVHHPESR	SSIQHSSVVP	SIATVVIIIS	840
VCMLVFVVM	GVYRVRIAHO	HFIQETEAAK	ESEMDWDDSA	LTITVNPMEK	HEGPGHGEDE	900
TEGEEEEEE	EEMSSSSGSD	DSEEEEEEEG	MGRGRHGQNG	ARQAQLEWDD	STLPY	

# BFG7 DNA sequence (SEQ ID NO:16)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Nucleic Acid Accession #: n/a; Coding sequence: <1-906 (stop codon underlined)

CGGGTCGACC	CACGCGTCCG	GGGAGAAAGG	ATGGCCGGCC	TGGCGGCGCG	GTTGGTCCTG	60
CTAGCTGGGG	CAGCGGCGCT	GGCGAGCGGC	TCCCAGGGCG	ACCGTGAGCC	GGTGTACCGC	120
GACTGCGTAC	TGCAGTGCAG	AGAGCAGAAC	TGCTCTGGGG	GCGCTCTGAA	TCACTCCCGC	180
TCCC GCCAGC	CAATCTACAT	GAGTCTAGCA	GGCTGGACCT	GTCGGGACGA	CTGTAAGTAT	240
GAGTGTATGT	GGGTACACGT	TGGGCTCTAC	CTCCAGGAAG	GTCACAAAGT	GCCTCAGTTC	300
CATGGCAAGT	GGCCCTTCTC	CCGGTTCCTG	TTCTTTCAAG	AGCCGGCATC	GGCCGTGGCC	360
TCGTTTCTCA	ATGGCCTGGC	CAGCCTGGTG	ATGCTCTGCC	GCTACCGCAC	CTTCGTGCCA	420
GCCTCCTCCC	CCATGTACCA	CACCTGTGTG	GCCTTCGCCT	GGGTGTCCCT	CAATGCATGG	480
TTCTGGTCCA	CAGTYTTCCA	CACCAAGGAC	ACTGACCTCA	CAGAGAAAAT	GGACTACTTC	540
TGTGCTTCCA	CTGTGCTCCT	ACACTCAATC	TACCTGTGCT	GCGTCAGCCT	CATCCGCTTC	600
GACTATGGCT	ACAACCTGGT	GGCCAACGTG	GCTATTGGCC	TGGTCAACGT	GGTGTGGTGG	660
CTGGCCTGGT	GCCTGTGGAA	CCAGCGGCGG	CTGCCCTCAC	TGCGCAAGTG	CGTGGTGGTG	720
GTCCTGTGTC	TGCAGGGGCT	GTCCCTGCTC	GAGCTGCTTG	ACTTCCCAAC	GCTCTTCTGG	780
GTCCTGGATG	CCCATGCCAT	CTGGCACATC	AGCACCATCC	CTGTCCACGT	CCTCTTTTTC	840
AGCTTTCTGG	AAGATGACAG	CCTGTACCTG	CTGAAGGAAT	CAGAGGACAA	GTTCAAGCTG	900

GACTGAAGAC	CTTGGAGCGA	GTCTGCCCCA	GTGGGGATCC	TGCCCCCGCC	CTGCTGGCCT	960
CCCTTCTCCC	CTCAACCCCT	GAGATGATT	TCTCTTTTCA	ACTTCTTGAA	CTTGGACATG	1020
AAGGATGTGG	GCCCAGAATC	ATGTGGCCAG	CCCACCCCT	GTTGGCCCTC	ACCAGCCTTG	1080
GAGTCTGTTC	TAGGGAAGGC	CTCCCAGCAT	CTGGGACTCG	AGAGTGGGCA	GCCCCCTCTAC	1140
CTCCTGGAGC	TGAACTGGGG	TGGAACCTGAG	TGTGCTCTTA	GCTCTACCGG	GAGGACAGCT	1200
GCCTGTTTCC	TCCCCATCAG	CCTCCTCCCC	ACATCCCCAG	CTGCCTGGCT	GGGTCTCTGAA	1260
GCCTCTGTGC	TACCTGGGAG	ACCAGGGACC	ACAGGCCTTA	GGGATACAGG	GGGTCCCCCT	1320
CTGTTACCAC	CCCCACCCCT	CCTCCAGGAC	ACCACTAGGT	GGTGTCTGGAT	GCTTGTCTCT	1380
TGGCCAGCCA	AGGTTACACG	CGATTCTCCC	CATGGGATCT	TGAGGGACCA	AGCTGCTGGG	1440
ATTGGGAAGG	AGTTTCACCC	TGACCRRTGC	CCTAGCCAGG	TTCCAGGAG	GCCTCACCAT	1500
ACTCCCTTTC	AGGGCCAGGG	CTCCAGCAAG	CCCAGGGCAA	GGATCCTGTG	CTGTGTGTCTG	1560
GTTGAGAGCC	TGCCACCGTG	TGTCGGGAGT	GTGGGCCAGG	CTGAGTGCAT	AGGTGACAGG	1620
GCCGTGAGCA	TGGGCCTGGG	TGTGTGTGAG	CTCAGGCACT	AGGTGCGCAG	TGTGGAGACG	1680
GGTGTGTGCG	GGGAAGAGGT	GTGCTTCAA	AGTGTGTGT	GTGCAGGGGG	TKGGTGTGTT	1740
AAGCGTGGGT	TAGGGGAACG	TGTGTGCGCG	TGCTGGTGGG	CATGTGAGAT	GAGTGACTGC	1800
CGGTGAATGT	GTCCACAGTT	GAGAGGTTGG	AGCAGGATGA	GGGAATCCTG	TCACCATCAA	1860
TAATCACTTG	TGGAGCGCCA	CTTGGCCCAA	GACGCCACCT	GGGCGGACAG	CAGGAGCTCT	1920
CCATGGCCAG	GCTGCTGTG	TGCATGTTCC	CTGTCTGGTG	CCCCTTTGCC	CGCCTCCTGC	1980
AAACCTCAC	GGGTCCCCAC	ACAACAGTGC	CCTCCAGAAG	CAGCCCTCG	GAGGCAGAGG	2040
AAGGAAAATG	GGGATGGCTG	GGGCTCTCTC	CATCCTCCTT	TTCTCCTTGC	CTTCGCATGG	2100
CTGGCCTTCC	CCTCAAAAC	CTCCATTCCC	CTGCTGCCAG	CCCCTTTGCC	ATAGCCTGAT	2160
TTTGGGGAGG	AGGAAGGGGC	GATTTGAGGG	AGAAGGGGAG	AAAGCTTATG	GCTGGGTCTG	2220
GTTCCTTCCC	TTCCACAGAG	GTCTTACTGT	TCCAGGGTGG	CCCCAGGGCA	GGCAGGGGCC	2280
ACACTATGCC	TGCGCCCTGG	TAAAGGTGAC	CCCTGCCATT	TACCAGCAGC	CCTGGCATGT	2340
TCCTGCCCCA	CAGGAATAGA	ATGGAGGGAG	CTCCAGAAAC	TTTCCATCCC	AAAGGCAGTC	2400
TCCGTGTTG	AAGCAGACTG	GATTTTGTCT	CTGCCCTGA	CCCCTTGTC	CTCTTGAGG	2460
GAGGGGAGCT	ATGCTAGGAC	TCCAACCTCA	GGGACTCGGG	TGGCTGCGC	TAGCTTCTTT	2520
TGATACTGAA	AACTTTTAAG	GTGGGAGGGT	GGCAAGGGAT	GTGCTTAATA	AATCAATTCC	2580
AAGCCTCAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAA			

# BFG7 Protein sequence (SEQ ID NO:17)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 117-133, 179-195, 211-227, 235-251, 266-282, 296-312; PFAM domains: none; Summary: A type III membrane protein of unknown function; is adjacent to HER2 on the genome, and its overexpression in breast cancer is highly correlated with HER2 expression; may be used to predict HER2 overexpression and amplification.

RVDPRVRGER	MAGLAARLVL	LAGAAALASG	SQGDREPVYR	DCVLQCEEQN	CSGGALNHFR	60
SRQPIYMSLA	GWTCRRDDCKY	ECMWVTVGLY	LQEGHKVPQF	HGKWPFPSRFL	FFQEPASAVA	120
SFLNGLASLV	MLCRYRTFVP	ASSPMYHTCV	AFAWVSLNAW	FWSTVFHTRD	TDLTEKMDYF	180
CASTVILHSI	YLCCVRTVGL	QHPAVVSAFR	ALLLLMLTVH	VSYLSLIRFD	YGYNLVANVA	240
IGLVNVVWWL	AWCLWNQRL	PHVRKCVVVV	LLLQGLSLE	LLDFPPLFWV	LDAHAIWHIS	300
TIPVHVLFFS	FLEDDSLYLL	KESEDKFKLD				

# BCN4 DNA sequence (SEQ ID NO:18)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Nucleic Acid Accession #: n/a; Coding sequence: 143-874 (start and stop codons underlined)

GGGAGGGAGA	GAGGCGCGCG	GGTGAAAGGC	GCATTGATGC	AGCCTGCGGC	GGCCTCGGAG	60
CGCGGCGGAG	CCAGACGCTG	ACCAGTTTCC	TCTCCTCGGT	CTCCTCCGCC	TCCAGCTCCG	120
CGCTGCCCCG	CAGCCGGGAG	CCATGCGACC	CCAGGGCCCC	GCCGCCTCCC	CGCAGCGGCT	180
CCGCGGCCTC	CTGTGCTCC	TGCTGCTGCA	GCTGCCCGCG	CCGTGAGCG	CCTCTGAGAT	240
CCCCAAGGGG	AAGCAAAGG	CGCAGCTCCG	GCAGAGGGAG	GTGGTGGACC	TGTATAATGG	300
AATGTGCTTA	CAAGGCGCAG	CAGGAGTGCC	TGGTCGAGAC	GGGAGCCCTG	GGGCCAATGG	360
CATTCCGGGT	ACACCTGGGA	TCCCAGGTCG	GGATGGATTC	AAAGGAGAAA	AGGGGGAATG	420
TCTGAGGGAA	AGCTTTGAGG	AGTCTGGAC	ACCAACTAC	AAGCAGTGTT	CATGGGATTC	480
ATTGAATTAT	GGCATAGATC	TTGGGAAAAT	TGCGGAGTGT	ACATTACAA	AGATGCGTTC	540
AAATAGTGCT	CTAAGAGTTT	TGTTCACTGG	CTCACTTCGG	CTAAAATGCA	GAAATGCATG	600
CTGTGAGCGT	TGGTATTTCA	CATTCAATGG	AGCTGAATGT	TCAGGACCTC	TTCCCATTGA	660
AGCTATAATT	TATTTGGACC	AAGGAAGCCC	TGAAATGAAT	TCAACAATTA	ATATTTCATG	720
CACCTTCTCT	GTGGAAGGAC	TTTGTGAAGG	AATTGGTGCT	GGATTAGTGG	ATGTTGCTAT	780
CTGGGTTGGC	ACTTGTTCAG	ATTACCCAAA	AGGAGATGCT	TCTACTGGAT	GGAATTCAGT	840
TTCTCGCATC	ATTATTGAAG	AACTACCAAA	ATAAATGCTT	TAATTTTCAT	TTGCTACCTC	900
TTTTTTTATT	ATGCTTGGGA	ATGGTTCAC	TAAATGACAT	TTTAAATAAG	TTTATGTATA	960
CATCTGAATG	AAAAGCAAAG	CTAAATATGT	TTACAGACCA	AAGTGTGATT	TCACACTGTT	1020
TTTAAATCTA	GCAATATTCA	TTTTGCTTCA	ATCAAAAGTG	GTTTCAATAT	TTTTTTTGTG	1080
TGGTTAGAAT	ACTTTCTTCA	TAGTCACATT	CTCTCAACCT	ATAATTGGA	ATATTGTTGT	1140
GGTCTTTTGT	TTTTTCTCTT	AGTATAGCAT	TTTTTAAAAA	ATATAAAGC	TACCAATCTT	1200
TGTACAATTT	GTAAATGTTA	AGAATTTTTT	TTATATCTGT	TAAATAAAAA	TTATTTCCAA	1260
CAACCTTAAA	AAAAAAAAAA	AAAAAA				

# BCN4 Protein sequence (SEQ ID NO:19)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Protein Accession #: n/a; Predicted Signal sequence: 1-30; TM domains: none; PFAM domains: none; Summary: a secreted protein; has a mouse orthologue (see sequence below).

MRPQGPAASP	QRLRGLLLLL	LLQLPAPSSA	SEIPKGKQKA	QLRQREVVDL	YNGMCLQGPA	60
GVPGRDGSPG	ANGIPCTPGI	PGRDGFKGEK	GECLRESFEE	SWTPNYKQCS	WSSLNYGIDL	120
GKIAECTFTK	MRSNSALRVL	FSGSLRLKCR	NACCQRWYFT	FNGAECGSL	PIEAIYLDQ	180
GSPMNSTIN	IHRTSSVEGL	CEGIGAGLVD	VAIWVGTCSD	YPKGDASTGW	NSVSRIIEE	240

LPK

Mouse BCN4 Protein sequence (SEQ ID NO:20)  
Gene name: ESTs; Unigene number: Mm.41556

XXXXAAPPQL	LLGLFLVLLL	LLQLSAPSSA	SENPVKKQKA	LIRQREVVDL	YNGMCLQGPA	60
GVPGRDGSPG	ANGIPGTPGI	PCQDGFKEK	GECLRESFEE	SWTPNYKQCS	WSSLNYGIDL	120
GKIAECTFTK	MRSNSALRVL	FSGSLRLYCR	NACCQRWYFT	FNGAECGPP	PIEAIXXXXX	180
XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXSD	YPKGDAYTGW	DSVSRIIEE	240

LPK